山形大学紀要（工学）第23巻　第2号　平成7年1月
Bull. Yamagata Univ. (Eng.) Vol.23 No. 2 Jan. 1995

A Concept of Consensus Genic Sequence of tRNA Genes from Bacteria as a Key to Clarify Genetic Mechanisms

KUDO Yoshihiro, KANAYA Shigehiko and KONNO Tomiharu

Department of Electrical and Information Engineering, Faculty of Engineering

(*Received September* 1, 1994)

Abstract

The first optimistic attempt was carried out to demonstrate usefulness of a concept of consensus genic sequence of tRNA, such as $tRNA^{Ala}$-$tRNA^{Met}$-$tRNA^{Ile/Met}$-$tRNA^{Ser}$, from bacteria as a promising key to clarify the complicated genetic mechanisms such as transcription to RNAs and translation to proteins. First of all, bacteria were easily classified into five categories, Mycoplasmas, Gram-positive bacteria, Gram-negative bacteria, Archaebacteria and Bacteria of unknown categories. Then among genic sequences of tRNA and rRNA genes from bacteria, common genic sequences were intuitively searched for. The next, from the resultant common sequences, consensus sequences were subjectively chosen. Unexpectedly and luckily, resultant consensus sequences proved to be clearly related to their proper categories of bacteria without special intention. On the basis of the fact, categories of some of bacteria in Bacteria of unknown categories were on trial.

1. Introduction

Besides the Pribnow box in promoter sites for transcription initiation, and the Shine-Dalgarno sequence associated with ribosome-binding sites for translation initiation, many kinds of consensus sequences play respective important roles on various stages in complicated genetic systems on various organisms. Nevertheless, the concept of"consensus genic sequence"of tRNA genes would have not been valued, in contrast with active studies on (I) DNA sequences of tRNA and rRNA genes and intergenic spacers flanking them, (II) genic constitution of clusters of tRNA genes, and (III) frequencies of anticodons, and though sometimes consensus-like genic sequences have been mentioned: e. g., a 4-gene partial sequence $tRNA^{Ala}$-$tRNA^{Met}$-$tRNA^{Ile}$-$tRNA^{Ser}$ in genomes from *Bacillus subtilis* [1),2)]and four *Mycoplasmas* species:*Acholeoplasma* laidlawii[3)], Mycoplasma mycoides[4)], *Mycoplasma capricolum*[5)], and *Spiroplasma melliferum*[6)]. We expect that consensus genic sequences of tRNA genes is promising keys to clarify a part of the complicated genetic mechanisms. Therefore, to our knowledge, this is the first optimistic and positive attempt to demonstrate the value of consensus genic sequences of tRNA.

2. Methodology

2. 1 Notation

Liner notations

Amino acid (anticodon)

and

Amino acid

are preferentially used as substituents of $tRNA^{\text{Amino acid}}_{\text{anticodon}}$ (or Amino $acid^{\text{anticodon}}$) and $tRNA^{\text{Amino acid}}$, respectively, where amino acid except methionine is described in the IUPAC/IUB one-character-code system. For example, W(CCA) is equivalent to $tRNA^{\text{Trp}}_{\text{CCA}}$ and Trp^{CCA}. Methionine is represented with either "M" or "X" which always indicates an initiator methionine.

5 S, 7 S,16S, and 23S rRNA genes are simply represented with 5, 7, 16 and 23, respectively, only in a genic sequence. For example, 16 S rRNA-$tRNA^{\text{Ile}}$-$tRNA^{\text{Ala}}$-23S rRNA-5 S rRNA-$tRNA^{\text{fMet}}$ will become 16 I A 23 5 X.

2. 2 Method

(1) First of all, bacteria are classified into some categories including "Bacteria of unknown categories" if necessary.

(2) As a preliminary, factors which may change the order of tRNA genes in a genic sequence are examined.

(3) Among genic sequences of tRNA and/or rRNA genes from bacteria, common genic sequences are intuitively searched for.

(4) From the resultant common sequences, candidates of consensus sequences are subjectively chosen and related to the categories classified in (1).

(5) Bacteria of unknown categories are classified into the categories on trial.

3. Result

3. 1 Classification of bacteria

Bacteria were easily classified into five categories:

Mycoplasmas(M),

Gram-positive bacteria(P),

Gram-negative bacteria(N),

Archaebacteria(A),

and Bacteria of unknown categories(U).

A bacterial species is given a unique code consisting of category code and a sequential number as follows: e.g.,

Acholeplasma laidlawii	M 1
Bacillus subtilis	P 1
Escherichia coli	N 1
Thermofilum pendens	A10
and *Rhodobacter sphaeroides*	U 3.

3. 2 Data variation

Among various kinds of data variation, only the difference between genic sequences to be the same of tRNA has interest for the present paper.

Case 1. Green et al[8] showed a DNA sequence containing the portion of a 9-tRNA gene cluster, 5 V T K L G L R P A 16, from *B. subtilis*.

The GenBank registers this(Locus BACTG 9168), where the value indicating the position of "T" of the "5 V T K……" is different from the source and means a strained order (5 T V K……):

	The source	GenBank
5	1..72	as it is
V	95..170	as it is
T	175..250	75..250
K	287..362	as it is

Interpreting "75..250" (abnormally long!) as a clerical error, the present paper adopts "175.. 250".

Case 2. Green et al[2] determined an entire contiguous nucleotide sequence including a 21-gene sequence of tRNA from *B. subtilis* by connecting three short sequences with their own experimental result. According to the authors,

A. The DNA sequence and its associated sequences

	Original Pos. Nucl. P		Green 1985 Pos. Nucl. Q		GenBank BACRGRRNB R	Original Pos. Nucl. S		Remarks Q-R	Q-P	Q-S
	(Stewart 1983)					& GenBank BACTGTRNE				
	1-5	no data	1-5	no data					0	
1	578	G	578-9	CG	573-4			5	0	
	1172	end	1173	end	1168			5	1	
	(Green 1983)					(Wawrousek 1984)				
2	135	A	5035-6	GA	5030-1			5	4901	
3	253	C	5154	C	5149	47	T	5	4901	5107
4	423	T	5324	T	5319	217-8	GT	5	4901	5106
5	844	A	5745	A	5740	639	T	5	4901	5106
6	1035	A	5986	A	5981	880-1	AA	5	4901	5105
7	1270	T	6171-2	CT	6161-2	1066	T	5	4902	5105
8	1322	T	6224	T	6219	1118-9	TT	5	4902	5106
9	1485	G	6387	G	6382	1282-3	GG	5	4902	5104
10	1633	T	6535	T	6530	1431-3	CAT	5	4902	5102
11	2345-6	CC	7247-8	CC	7242-3	2145	C	5	4902	5103
12	2351	A	7253	A	7248	2150-1	GC	5	4902	5102
13	2395-6	TA	7297-8	TA	7292-3	2195	A	5	4902	5103
14	2484	T	7386	T	7381	2283	G	5	4902	5103
15	2499-0	CG	7401-2	CG	7496-7	2298-9	GC	5	4902	5103
16	2506-7	GG	7408-9	GG	7403-4	2305	G	5	4902	5104
	2533	end	7435	end	7430	(2331	A)	5	4902	5104
						2497	end			

the relation between the sequences above and below

B. The genic sequence.

1	2	3	4		5		6		7	8		9	10						11-16
16	23	5	V	T	K	L	G	L	R	P	A	M	*	S	X	D	F	H	G	I	N	S	E

I (Green 1983, 1985)
M (Wawrousek 1984)

Figure1. Data variation in DNA sequences of 21-gene genic sequence of tRNA from Batillus subtilis.

"The bp 1-1170 are from Stewart and Bott (1983), 1813-2275 (with minor correction) from Bott and Hollis(1982), and 4901- end from Green and Vold (1983) (confirmed by Wawrousek et al.,1984)."

Figure 1 shows sixteen different sites

between the three original sequences and the resultant entire one.

Among them, the present paper is affected by assignment of the second $tRNA^{CAT}{}_S$(which are the same even on the level of DNA sequence) to different amino acids: I (CAT) [1),2)]and M (CAT)[9)] because they lead to different genic sequences, ……A M I S…… and ……A M M S……, respectively. In fact, they are exactly the same as shown in Figure

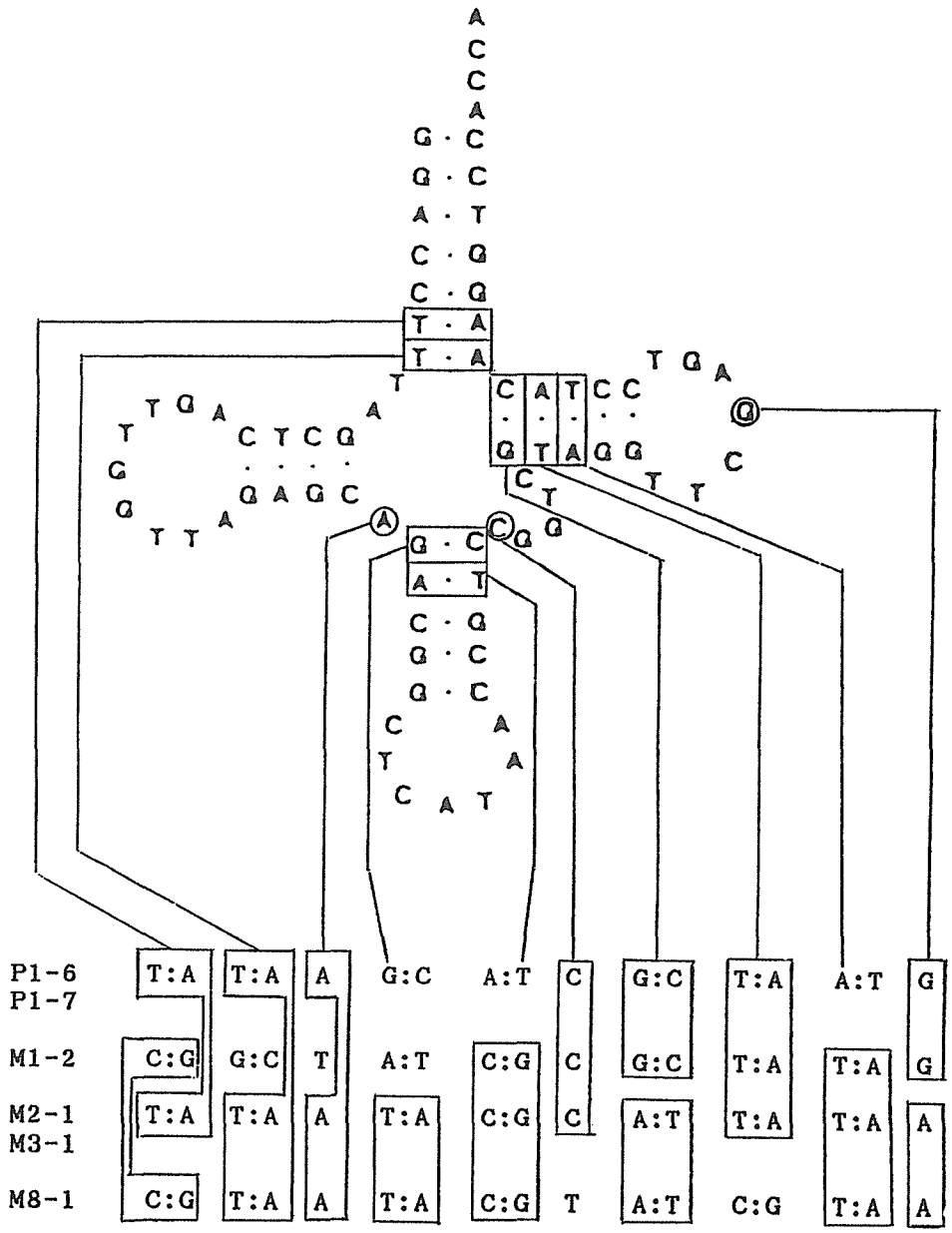

Figure 2. Variation among the second $tRNA^{CAT}$ sequences of the six genic sequences containing the 4-gene partial genic sequence, A M M S (M3-1) or A M I S (the others).

2 (P1-6 and P1-7). Because of no experimental evidence, ingnoring this difference, the present paper assumes that they are equivalent to each other.

3. 3 Data of genic sequences of tRNA and rRNA genes from bacteria

All of necessary sequences were extracted from the GenBank data which were provided as Distribution No.16(Jan.,1994) by DNA Data Base of Japan (DDBJ)[7]. Table 1 lists the

Table1. Genic sequences of rRNA and tRNA from bacteria.

The five categories OF bacteria		
1. Mycoplasmas (M)		
2. Gram positive bacteria (P)		
3. Gram negative bacteria (N)		
4. Archeobacteria (A)		
5. Bacteria of unknown categories (U)		
Mycoplasmas (M)		
Acholeplasma laidlawii (M1)		16 23
	M1-1	16 I A 23
	M1-2	23 5 V T K L A M I S M D F
	M1-3	S E
	M1-4	H Q L
	M1-5	23 5 N
Mycoplasma capricolum (M2)		16 23
	M2-1	N E V T
	M2-2	R P A M I S X D F
	M2-3	K L
	M2-4	K L 16
	M2-5	T Y Q K L
	M2-6	W W
Mycoplasma mycoides (M3)	M3-1	N E V T
	M3-2	R P A M I S X D F
Mycoplasma PG50 (M4)		16 23
	M4-1	K L 16
Mycoplasma pneumoniae (M5)	M5-1	T Q K L G
Mycoplasma arginini, Mycoplasma arthritidis, Mycoplasma fermentans, Mycoplasma gallisepticum, Mycoplasma hominis, Mycoplasma hyopneumoniae, Mycoplasma hyorhinis, Mycoplasma neurolyticum, Mycoplasma orale, Mycoplasma pulmonis, Mycoplasma salivarium, Mycoplasma F38, and Ureaplasma urealyticum		16 23
Mycoplasma-like sp. (M6)		16 23, 23 5
	M6-1	Y 16 I 23
Spiroplasma citri (M7)		16 23
	M7-1	W W S
Spiroplasma melliferum (M8)	M8-1	C R P A M I S X D F

(Cont,d)

(Table 1. Cont,d)

Gram positive bacteria (P)		
Bacillus subtilis (P1)		23 5, 5 16
	P1-1	K E D F
	P1-2	16 I A 23
	P1-3	5 V T K L G L R P A 16
	P1-4	23 5 M D
	P1-5	23 5 N T G R P A 16
	P1-6	16 23 5 V T K L G L R P A M I S X D F H G I N S E
	P1-7	23 5 V T K L G L R P A M M S X D F H G I N S E
	P1-8	23 5 N S E V M D F T Y W H Q G C L L
Bacillus PS3 (P2)	P2-1	N S E V M D
Chlostridium botulinum (P3)		23 5
Chlostridium perfringens, Desulfurococcus mobilis, Listeria monocytogenes, Mycobacterium bovis, Mycobacterium tuberculosis, and Streptomyces coelicolor		16 23
Chlostridium tyrobutyricum, Frankia (AcN14a), Frankia (ORS020606), Mycobacterium leprae, and Streptomyces griseus		16 23 5
Enterococcus hirae		r r, r t r
Lactobacillus delbrueckii (P3)		
	P3-1	N P G R V D
	P3-2	5 G R V Q D
Lactococcus lactis (P4)	P4-1	16 A 23 5 N
Micrococcus luteus (P5)	P5-1	G C V G
	P5-2	T M
Mycobacterium smegmatis (P6)	P6-1	G P
Staphylococcus aureus (P7)	P7-1	X D F
Staphylococcus hominis (P8)	P8-1	V M D F
Staphylococcus warneri (P9)	P9-1	X D F
Storeptococcus pneumoniae (P10)	P10-1	16 A 23
Storeptococcus pyogenes (P11)	P11-1	X F
Storeptococcus salivarius (P12)	P12-1	23 5 N
Storeptococcus sp. group B (P13)	P13-1	X F
Storeptococcus sp. group G (P14)	P14-1	X F
Streptomyces lividanlls (P15)		23 5, 16 23
	P15-1	G C V V V
	P15-2	S R
	P15-3	N N
	P15-4	E Q E E Q E

(Table 1. Cont,d)

Gram negative bacteria (N)

Escherichia coli (N1)		23 5
	N1-1	5 D W
	N1-2	5 T 5
	N1-3	16 E 23 5
	N1-4	16 E 23
	N1-5	16 I
	N1-6	16 I A 23
	N1-7	16 I A 23 5
	N1-8	23 5 T 5
	N1-9	23 5 D
	N1-10	23 5 D W
	N1-11	A A
	N1-12	R H L P
	N1-13	D W
	N1-14	G C L
	N1-15	G G G
	N1-16	I A 23
	N1-17	L G M Q
	N1-18	L L L
	N1-19	K V K
	N1-20	S R R R
	N1-21	T Y
	N1-22	T Y G T
	N1-23	Y G T
	N1-24	M M
	N1-25	V V K
	N1-26	V V V K
	N1-27	V V
	N1-28	< V V V K, A A
	N1-29	<W E, 16 I A 23
Aeromonas hydrophil (N2)	N2-1	R H L P
Campylobacter jejuni (N3)	N3-1	16 A L
Caulobacter crescentus (N4)	N4-1	16 I
	N4-2	16 I A 23
Chlamydia trachomatis, and Proteus vulgaris		16
Photobacterium leiognathi (N5)	N5-1	M L
Photobacterium phosphoreum (N6)	N6-1	R H
	N6-2	P H P
	N6-3	P P P H P P H P
Plesiomonas shigelloides (N7)	N7-1	16 E 23
Psudomonas aeruginosa (N8)	N8-1	I A
	N8-2	V D D
	N8-3	Y G T
Salmonella typhimurium (N9)	N9-1	R H L P
Sulfolobus solfataricus (N10)	N10-1	M V
	N10-2	H G
	N10-3	S L

(Table 1. Cont,d)

Thiobacillus ferrooxidans (N11)		N11-1	16 I A 23
Vibrio harveyi (N12)		N12-1	R H P
Yersinia enterocolitica			23
Archeobacteria (A)			
Halobacterium cutirubrum	(A1)	A1-1	23 5 C
Halobacterium halobium	(A2)	A2-1	16 A 23 5
Halobacterium marismortui	(A3)	A3-1	S L
Halobacterium volcanii	(A4)	A4-1	5 C
		A4-2	W W
Metanobacterium thermoautotrophieum	(A5)	A5-1	7 S 16 A 23 5
Methanococcus vannielii	(A6)		16 23, 23 5
		A6-1	16 A 23
		A6-2	T P Y K 5 D K D
Methanosarcina frisius	(A7)	A7-1	A 16 23
Methanothermus fervidus	(A8)	A8-1	7 S 16 A
		A8-2	N M E L H
		A8-3	T P D K
Methanothrix soehngenii	(A9)	A9-1	16 A 23
Thermococcous celer	(A10)	A10-1	5 D
Thermofilum pendens	(A11)	A11-1	16 23 M, <G
Thermoproteus tenax			A, <A, M, L
Bacteria of unknown categories (U)			
Anacystis nidulans	(U1)		23 5
		U1-1	16 I A
		U1-2	16 I A 23
		U1-3	A 23 5
Borrelia burgdorferi			5 23
Haemophilus influenzae	(U2)	U2-1	G L K
Rhodobacter sphaeroides	(U3)	U3-1	16 I A 23 5 X
Thermotoga maritima	(U4)	U4-1	M M T Y R
Thermus thermophilus	(U5)	U5-1	T Y G T

bacteria concerned, and genic sequences of tRNA and/or rRNA genes from the bacteria, of which, all of sequences containing tRNA genes were given their respective identifiers. For example, M1-2 indicates the second sequence from M1 (A. *laidlawii*).

3. 4 Consensus RNA-genic sequences

Grouping of similar genic sequences in Table 1 was carried out in a subjective and arbitrary but reasonable way. For example, of two "W W" sequences, only M2-6, was enclosed together with M7-1 (W W S) in group E (W W S), and A4-2 was not. Neither N1-20 (S R R R) and P15-2 (S R) [group R(a)] nor N10-3 (S L) and A3-1 (S L) [group R(b)] were grouped with. P10-1 (16 A 23) was not connected with 7

Table 2. Common and consensus genic sequences.

```
 * .   A putative  consensus sequence
===   separates different categories
-----------------------------------------------------------------

A
  *                      K L ---------------------------------16
  *         23 5 V T   Q K L     R P A       X D F
  *         23 5 N       K L G L     A M I S       H G I N S E
  ---------------------------------------------------------------
   M2-5             T Y Q K L
   M5-1             T   Q K L G
   M2-4                   K L                                  16
   M4-1                   K L                                  16
   M2-3                   K L
   P1-6 16 23 5 V T       K L G L R P A M I S X D F H G I N S E
   P1-7    23 5 V T       K L G L R P A M M S X D F H G I N S E
   P1-3       5 V T       K L G L R P A                        16
   M1-2    23 5 V T       K     L     A M I S M D F
   M8-1  C                          R P A M I S X D F
   M2-2                             R P A M I S X D F
   M3-2                             R P A M I S X D F
   P1-5    23 5 N T           G     R P A                      16
   P7-1                                     X D F
   P9-1                                     X D F
   P1-4    23 5                         M     D


B (#  See Group N)
 *              23 5 N S E V D M F T Y W H Q
 *        16 A 23 5 N
 ---------------------------------------------------------------
  M1-3                  S E
  M1-4                                   H Q     L
  M1-5           23 5 N
 (P1-5           23 5 N T                     G          R P A 16)
  P4-1     16 A 23 5 N
  P10-1    16 A 23
  P12-1          23 5 N
  P1-8           23 5 N S E V M D F T Y W H Q G C L L
  P2-1                N S E V M D
  P8-1                    V M D F
 ===============================================================
  N1-18                                              L L L

 #A2-1     16 A 23 5
 #A6-1     16 A 23
 #A9-1     16 A 23
```

(Cont,d)

(Table 2. Cont,d)

C

*	R H
*	L P
N1-12	R H L P
N2-1	R H L P
N6-1	R H
N9-1	R H L P
N12-1	R H P

D

*	X F
P11-1	X F
P13-1	X F
P14-1	X F

E

*	W W S
M7-1	W W S
M2-6	W W
A4-2	W W

F

*	N E V T
M2-1	N E V T
M3-1	N E V T

G

*	T P 5 D
*	D K
A6-2	T P Y K 5 D K D
A8-3	T P D K
A10-1	5 D

H

*	16 I A 23
M1-1	16 I A 23
P1-2	16 I A 23

H or I

U1-1	16 I A
U1-2	16 I A 23
U1-3	A 23

I

*	16 I A 23 5 X
U3-1	16 I A 23 5 X
N1-7	16 I A 23 5
N1-5	16 I
N1-6	16 I A 23
N1-16	I A 23
N1-29	16 I A 23
N4-1	16 I
N4-2	16 I A 23
N11-1	16 I A 23
N8-1	I A
*	23 5 T 5
N1-8	23 5 T 5
N1-2	5 T 5
*	23 5 D W
N1-1	5 D W
N1-9	23 5 D
N1-10	23 5 D W
N1-13	D W
*	16 E 23 5
N1-3	16 E 23 5
N1-4	16 E 23
N7-1	16 E 23

J

*	A 16 23 M
A7-1	A 16 23
A11-1	16 23 M

K

*	V V V K
*	V K
N1-19	K V K
N1-25	V V K
N1-26	V V V K
N1-27	V V
N1-28	V V V K

(Table 2. Cont,d)

L

*				G	R	V		
P3-1		N	P	G	R	V		D
P3-2	5			G	R	V	Q	D

M

*			T	Y	G	T
*		M	M	T	Y	
N1-24	M	M				
N1-21			T	Y		
N1-22			T	Y	G	T
N1-23				Y	G	T
N8-3				Y	G	T
U4-1	M	M	T	Y		R
U5-1			T	Y	G	T

N (# See also Group B)

*	7	S	16	A	23	5	C
A5-1	7	S	16	A	23	5	
A8-1	7	S	16	A			
#A6-1			16	A	23		
#A9-1			16	A	23		
#A2-1			16	A	23	5	
A1-1					23	5	C
A4-1						5	C

O

*	G	C	V
P15-1	G	C	V V V
P5-1	G	C	V G

P

*	P	P	P	H	P	P	H	P
N6-3	P	P	P	H	P	P	H	P
N6-2			P	H	P			

Q

*	A	A
N1-11	A	A
N1-28	A	A

R (Not grouped as yet)

(a)

N1-20	S	R	R	R
P15-2	S	R		

(b)

N10-3	S	L
A3-1	S	L

(c)

M6-1	Y	16	I	23	
P1-1	K	E	D	F	
P5-2	T	M			
P6-1	G	P			
P15-3	N	N			
P15-4	Q	E	E	Q	E
N1-14	G	C	L		
N1-15	G	G	G		
N1-17	L	G	M	Q	
N1-18	L	L	L		
(See also group B)					
N1-29	W	E			
N3-1	16	A	L		
N5-1	M	L			
N8-2	V	D	D		
N10-1	M	V			
N10-2	H	G			
A4-2	W	W			
(See also group E)					
A8-2	N	M	E	L	H
U2-1	G	L	K		

Table3. Common sequences as candidates of consensus sequences.

Group	Genic sequence	Categories of bacteria (Individual)	Categories of bacteria (Whole)	Group
A	23 5 V T	P1,M1		
	Q K L	M2,M5		
	K L 16	M2,M4		
	G L R P A	P1		
	R P A	M2,M3,M8,P1		
	A M I S	M1,M2,M3,M8,P1		
	X D F	M1,M2,M3,M8,P1,P7,P9		
	K L	M2,M4,M5,P1		
	H G I N S E	P1,P2	M&P	A
B	23 5 N S E V D M F T Y W H Q	M1,M12,P4,P8,P1,P2		
	16 A 23 5 N	P1,P4,P10	M&P	B
C	L P	N1,N2,N9		
	R H	N1,N2,N9,N6,N12	N	C
D	X F	P11,P13,P14	P	D
E	W W S	(M2,)M7	M	E
F	N E V T	M2,M3	M	F
G	T P	A6,A8		
	5 D	A6,A10		
	D K	A6,A8	A	G
H	16 I A 23	M1,P1 (U1 ?)	M&P	H
I	16 I A 23 5 X	U3(,N1,N4,N8,N11) (U1 ?)		
	23 5 T 5	N1		
	23 5 D W	N1		
	16 E 23 5	N1,N7	N	I
J	A 16 23 M	(A7,A11)	A	J
K	V V V K	N1	N	L
L	G R V	P3	P	K
M	T Y G T	U5,N1(,N8)		
	M M T Y	U4(,N1)	N	M
N	7 S 16 A 23 5	A5(,A1,A2,A7,A8,A9)		
	23 5 C	A1(,A4)	A	N
O	G C V	P5,P15	P	O
P	P P P H P P H P	N6	N	P
Q	A A	N1	N	Q

S 16 A 23 5 C(Group N). On the other hand, fusion (M&P) of the categories M (Mycoplasmas) and P(Gram-positive bacteria) was allowed whenever necessary. Table 2 shows resultant groups, A to Q.

Table 3 highlights the relation to the categories of bacteria.

4. Consideration

Table 3 suggests some direct and simple relation between groups of the genic sequences and the four categories (M, P, N, and A) except a couple of categories M and P. This would be a lucky accident beyond expectation because all the methods adopted (i. e., those for classification of bacteria, sorting of genic sequences extracted from the GenBank, and choice of consensus genic sequences) are not special but common ones, and were used without any special intention. That is, the flow of the procedure was very straight forward and natural.

The case of group E discriminatively treated two sequences, M 2-5 and A 4-2, of "W W". Nevertheless, it should be emphasized also that it was not a special technique that a potential insider (resp. outsider) of a group was treated from an optimistic (resp. a pessimistic) viewpoint. Thus it can be considered (i) that M 2-6

Table 4. Corresponding tRNA genes in the six sequences containing the 4-gene sequence A M I/M S (See also Table 1)

tRNA	Anticodon								
C	GCA								M8-1
V	TAC	P1-6	=	P1-7	M1-2				
T	TGT	P1-6	=	P1-7	M1-2				
K	TTT	P1-6	=	P1-7	M1-2				
L	CAG	P1-6	=	P1-7					
G	GCC	P1-6	&	P1-7					
L	TAA	P1-6	=	P1-7	M1-2				
R	ACC	P1-6	=	P1-7		M2-2	=	M3-2	M8-1
P	TCG	P1-6	=	P1-7		M2-2	=	M3-2	M8-1
A	TGC	P1-6	=	P1-7	M1-2	M2-2	=	M3-2	M8-1
M	CAT	P1-6	#	P1-7	M1-2	M2-2		M3-2	M8-1
I/M	CAT	P1-6	=	P1-7	M1-2	M2-2	=	M3-2	M8-1
S	TGA	P1-6	#	P1-7	M1-2	M2-2	#	M3-2	M8-1
X	CAT	P1-6	=	P1-7	MI-2	M2-2	&	M3-2	M8-1
D	GTC	P1-6	=	P1-7	M1-2	M2-2	&	M3-2	M8-1
F	GAA	P1-6	=	P1-7	M1-2	M2-2	=	M3-2	M8-1
H	GTC	P1-6	*	P1-7					
G	TCC	P1-6	=	P1-7					
I	GAT	P1-6	=	P1-7					
N	GTT	P1-6	=	P1-7					
S	GCT	P1-6	=	P1-7					
E	TTC	P1-6	=	P1-7					

= : Exact match on the level of DNA sequence
: Extra arms are a little different.
* : P1-7 = P1-6 + (a nucleotide at the -1 position).
& : Only one pair of nucleotides is different.

is a part of "W W S" and A 4-2 not, (ii) that P15-2 (S R) is not a part of "S R R R" forming N 1-20, and (iii) that N 10-3 (S L) and A 3-1 (S L) are very short parts of two longer different sequences. This discriminative treatment is not only natural but also reasonable and necessary because all the sequences discussed are short fragments whose flanking genic seguences are unknown.

On the other hand, M&P, fusion of the categories M (Mycoplasmas) and P (Gram-positive bacteria), is allowed if necessary. *Mycoplasmas,* wall-less prokaryotes, are sophylogenetically close to Gram-positive eubacteria such as *Bacillus* spp. to have ever been classified in them.

The reasonability of M&P also would be indirectly supported not only by that the very characteristic 4-gene partial sequence, $tRNA^{Ala}$-$tRNA^{Met}$-$tRNA^{Ile/Met}$-$tRNA^{Ser}$, (i.e., A M I/M S) is found in the sequences of M&P, such as P 1-6, P 1-7, M 1-2, M 2-2, M 3-2 and M 8-1, but also by that it has been never found as yet from bacteria of other categories such as *Escherichia* coli (N 1). Corresponding tRNA genes of the six sequences are very similar to as exemplified by I/M (CAT) shown in Figure 2, and Table 4 shows how they are the same on the level of anticodons in all cases, and have even exactly the same DNA sequences in many cases.

In brief, our first optimistic attempt may, at least partly, succeed to search the consensus genic sequences for any important key to clarify genetic mechanisms. If it is true, U 3 (Group I), and U 4 and U 5 (Group M) may be classified into N (Gram-negative bacteria) as proposed in Table 3 on trial, although at least anticodons of the tRNA concerned have to be examined. U 1 has too poor information to be classified into M&P (Group H), N (Group I), or any others

5. Conclusion

Usefulness of the "concept of consensus genic sequence" of tRNA from bacteria as a tool to clarify genetic mechanisms was suggested by a fact that its clear relation to a simple set of categories of the bacteria was found.

References

1) Christopher J. Green, and Barbara S. Vold: Nucl. Acids Res.,11, 5763(1983).
2) Christopher J. Green, George C Stewart, Mary Ann Hollis, Barbara S. Vold, and Kenneth F. Bott: Gene,37, 261(1985).
3) Reiji Tanaka, Yoshiki Andnchi, and Akira Muto: Nucl. Acids Res.,19, 6787(1991).
4) Tore Samuelsson, Per Elias, Florentyna Lustig, and Youssef S. Guindy: Biochem. J.,232, 223(1985).
5) Akira Muto, Yoshiki Andachi, Harumi Yuzawa, Fumiaki Yamano, and Shozo Osawa:Nucl. Acids Res.,18, 5037(1990).
6) M. J. Rogers, A. A. Steinmetz and R. T. Walker: Nucl. Acids Res., 14, 3145 (1986).
7) Sanzo Miyazawa: Computers and DNA, the proceedings of the interface between computation science and nucleic acid sequencing workshop, held December 12 to 16, 1988 in Santa Fe, New Mexico, Eds. Geroge I. Bell and Thomas G. Marr, (Addison-Wesley), pp.47-61(1990).
8) Christopher J. Green, and Barbara S. Vold: J. Bacteriol.,174, 3147(1992).
9) Eric F. Waerousek, Nalini Narasimhan, and Norman Hansen: J. Biol. Chem.,259, 3694(1984).